DOCUMENT RESUME

ED 354 034 JC 930 074

AUTHOR Yao, Min; Souza, Tony
TITLE Developing and Validating an Intermediate English Reading Test for English-as-a-Second-Language Learners.
INSTITUTION Vancouver Community Coll., British Columbia.
PUB DATE 18 Dec 92
NOTE 27p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Community Colleges; *Diagnostic Tests; *Educational Testing; *English (Second Language); Foreign Countries; Language Proficiency; Language Skills; *Language Tests; Limited English Speaking; Second Language Programs; *Student Placement; Test Construction; *Test Validity; Two Year Colleges
IDENTIFIERS Intermediate English Reading Test; Vancouver Community College BC

ABSTRACT

In order to more accurately place its intermediate English-as-a-Second-Language students in English courses, Vancouver Community College at King Edward Campus (KEC) developed the Intermediate English Reading Test (IERT). The goal was to create a test of minimum length that would yield scores with the necessary degree of reliability and validity for fine grading people whose English command is at the intermediate level. After consulting with reading instructors in the English Language Skills Department, a 20-minute test was devised consisting of a 1-page reading passage, an 8-question vocabulary section, and a 7-question comprehension section. Pilot testing was completed over 2 days with 344 intermediate level students at KEC speaking 56 different languages as their first language. The evaluation showed that the two sections were at the appropriate difficulty level for testees, but the comprehension section proved to be a slightly more reliable measure than the vocabulary section. The relatively moderate reliability of the test is viewed as acceptable because two questions in the vocabulary section artificially lowered the reliability rating of the whole test. In addition, the IERT is only one subtest of a battery of intermediate English tests used at KEC. Generally, the IERT may be regarded as a reliable reading test for people whose English is at the intermediate level. Tables of score distributions by test item and test section, a list of references, and the IERT are included. (MAB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED354034

Developing and Validating An Intermediate English Reading Test For English-as-a-Second-Language Learners

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. Yao

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

December 18, 1992

Min Yao and Tony Souza
Vancouver Community College
1155 East Broadway
Vancouver, B.C.
V5N 5V1, Canada
Tel. (604) 871-7171
E-mail myao@sfu.ca

JC930074

BEST COPY AVAILABLE

ABSTRACT

The Intermediate English Reading Test (IERT) was created to meet a need for more accurate testing and placement of ESL students at the intermediate level at Vancouver Community College. This article is essentially in two parts. The first part gives the necessary background information for the test, the criteria upon which the test was created, the test contents, and the arrangements made for pilot testing the prototype in intermediate level classes. The second part deals with the evaluation of the test itself. This includes the various criteria employed in the interpretation of the test results, tables listing test data, and complete analyses of the test results and data to determine reliability and validity of the IERT.

Developing and Validating the Intermediate English Reading Test[1]

TEST BACKGROUND

The Assessment Center of Vancouver Community College at King Edward Campus (KEC) provides English-as-a-Second-Language (ESL) testing services to a number of educational and social institutions for the assessment and placement of the large numbers of immigrants and refugees wishing to enter schools in and around the Vancouver region.

Two types of ESL tests are being used by the Assessment Center; Desk Tests for lower-level students from the Beginners to Intermediate levels, and the English Language Assessment (ELA) test for Advanced- or higher-level students.[2]

All new students wishing to enter KEC are tested by the Assessment Center. Preliminary placement is made by a staff member based on the conversation he/she has with the student when the latter comes into the Assessment office to make enquiries. Based on this, the student is given either the Beginners or the Intermediate Desk Test.

Until recently, the Intermediate Desk Test had comprised 3 sections - grammar and paragraph writing, to be done in 30 minutes, plus a 10 minute oral interview. If students scored above a certain mark, they were sent on to do the ELA, which is a more comprehensive three-hour test covering all skills - speaking, listening, reading, writing and grammar. Over a long period of time, it had been noted that many students that were sent on from the Desk Test to the ELA tested out ultimately at the Upper Intermediate level. In other words, the Desk Test was allowing many Upper Intermediate level students to go through to the ELA. As a result, it became obvious that it was necessary to tighten up the Desk Test. The Intermediate English Reading Test (IERT), as an additional part of the Desk Test, was created in response to this requirement.

TEST CONSTRUCTION CRITERIA

Our rationale for developing the IERT was to arrive at test of minimum length that would yield scores with the necessary degree of reliability and validity for fine grading people whose English command is at the intermediate level.

The IERT was constructed with two basic criteria in mind:

a) Students should be able to complete the test within a maximum time limit of 20 minutes. This, together with the existing 40 minutes spent on grammar, writing and the oral interview, would add up to a total test time of one hour. The time period was arbitrarily set, but the intention was to keep the Desk Test as short as possible without sacrificing too much accuracy of placement. The Desk Test was necessarily made short in order that the Assessment Center might be able to process as many students as possible in the shortest time.

b) A four-choice format was utilized in order to facilitate scoring.

Furthermore, as for the test itself, other aspects were also taken into consideration, including the following:

a) The length of the reading passage had to be proportionate to the amount of time allotted for this section of the test.

b) The vocabulary and grammatical structures in the passage had to be at a level which intermediate level students can handle.

c) The vocabulary questions had to be centered around words whose meanings could be discerned from the context.

e) The comprehension section had to cover basic reading skills such as identifying main ideas, making inferences and reading for details.

TEST CONTENTS

The prototype IERT was then created by consulting reading instructors at the English Language Skills Department and consisted of 3 parts:

a) A reading passage, which is at the optimum length of one page (double-spaced), designed specifically for intermediate level students;

b) A vocabulary section, consisting of 8 questions based on selected words and expressions in the passage. Within the passage itself, the selected vocabulary is underlined to facilitate location by the students.

c) A comprehension section, consisting of 7 questions on identifying main ideas, making inferences, and reading for details.

It was felt that these fifteen items would best represent the test domain of intermediate English reading.

PILOT TEST ARRANGEMENTS

Once the prototype IERT was assembled, it had to be pilot tested for reliability. With the large ESL programme at KEC, there is a ready-made audience on whom we can test prototype ideas, projects, and techniques. We decided to pilot test the newly created IERT in the English Language Skills Department. First, the Intermediate Level Coordinator was consulted. She was asked to approach her instructors with the idea of gathering as many volunteer classes as possible. Ultimately, all the intermediate instructors volunteered their classes, with the result that we had 343 Upper and Lower Intermediate level testees (169 from low-intermediate English classes; 174 from upper-intermediate English classes). The demographical data indicated that the participating testees spoke a total of 56 languages as their first languages and came from all over the world. We felt that these testees were a fair sample in terms of culture and language.

Pilot testing was completed over a period of two days. Staff from the Assessment Center went to the various classes at a pre-arranged time. Students were informed either by their instructors or by the Assessment Center staff member about the testing project, so many were quite eager to assist us by doing the test. In return, we agreed to supply each class with a copy of

their results. After the test, the answer sheets were gathered up and put through the computerized scanner in the Assessment Office for scoring.

ITEM ANALYSIS

The criteria for our item analysis were based on the guidelines set by Ebel (1965) for interpretation of discrimination index:

1. If Discrimination Index is greater than or equals 0.4, the item is functioning quite satisfactorily.

2. If Discrimination Index falls between 0.3 and 0.39, little or no revision is required.

3. If Discrimination Index is between 0.2 and 0.29, the item is marginal and needs revision.

4. If Discrimination Index equals or smaller than 0.19, the item should be eliminated or completely revised.

The item difficulty level was also taken into consideration in the process of item analysis. To maximize the test variance so that the test would legitimately represent the domain knowledge tested (reading comprehension and vocabulary), the range of tolerance for item difficulty level was set from .3 through .8.

The results of the item analysis are satisfactory. All the 15 items meet the above criteria. That is, all the 15 items have acceptable discrimination power and are at the right difficulty level.

ESTIMATED RELIABILITY

Table 1 is a summary of the distributions of raw scores for each of the two parts and the total scores of the test. Skewness of distribution is an indicator of the degree of test scores being symmetrically distributed.[3] If a test is too easy, scores will pile up at the high-score end of the scale (negatively skewed), whereas when the test is too hard, scores will pile up at the low-score end of the scale (positively skewed). In the ideal situation where the score distribution is unskewed, we will have a skewness index of zero. As can be seen in Table 1, although neither of the two sections and the total test has an unskewed score distribution, all the distributions are only slightly skewed. This suggests that the two sections and the total test are at the appropriate difficulty level for the testees.

Kurtosis of distribution refers to the peakedness or flatness of score distributions.[4] A normal distribution, calculated with the formula given in the footnote, will yield a Kurtosis index of

0.263. If Kurtosis value is greater than 0.263, the distribution is flat (platykurtic); if less than 0.263, the distribution is peaked (leptokurtic). When a test score distribution takes the shape of a flat curve, it means that the testees' scores have similar frequencies all over the distribution range. When a test score distribution has the form of a sharp curve, it is the indication that testees' scores were restricted to a narrow range, a sign of the test having a poor discriminant function. Thus, we would like to avoid having a test score distribution with the shape of a sharp (leptokurtic) curve. The Kurtosis indices in Table 1 show that, while none of the two sections and the total test has a normal curve, they do not deviate too much from the normal distribution, confirming the results from the item analysis that the discriminant functions of these IERT items are fairly good.

The Mean Difficulty in Table 1 shows the average scores of each section as well as the total test. This is another indicator of a test's discriminant power. If the mean difficulty value is high, that is, most of the testees scored high, it means that the test is too easy and fails to separate high-ability testees from low-ability testees (ceiling effect). Similarly, when the mean difficulty value is low, it means that the test is too difficult and does not discriminate the testees well either (flooring effect). Table 1 shows that the mean difficulty values of the two sections and the total test are 0.48, 0.61 and 0.54, respectively.

Interestingly enough, the first section, Vocabulary, is a little more difficult than the second section, Comprehension. However, the total test difficulty is balanced at the appropriate level. With these mild difficulty levels, we are convinced that the Intermediate English Reading Test (IERT) is at the appropriate difficulty level for the target testees, i.e., people whose English is at the intermediate level.

TABLE I.

Descriptive data of the Intermediate English Reading Test from a sample of 343 testees

	Vocabulary	Comprehension	Total
Number of Items	8	7	15
Skewness of Distribution	-0.236	0.405	0.121
Kurtosis of Distribution	0.25	0.30	0.25
Mean Difficulty	0.48	0.61	0.54

Because it was not practical to administer the IERT to the same 343 testees twice, the reliability for the IERT was estimated by internal consistency. Three procedures of internal consistency estimation were used in this report: coefficient alpha (computing item variances), Guttman's coefficient (split-form

method), and Spearman-Brown's prophecy (lower bound estimate). All these procedures yield values that are functions of the correlation between separately scored parts of the IERT. What we were looking for was testees' consistent performance of items across the two sections within the IERT. Such consistent performance would give us the confidence that the testees will perform in the same manner with other possible items in the content domain, i.e., good reliability. Table 2 is a summary of the reliability estimation of the two sections as well as the total test.

TABLE II.
Reliability estimation of
the Intermediate English Reading Test

	Vocabulary	Comprehension	Total
Guttman Split-Half	0.29	0.44	0.48
Spearman-Brown Prophecy	0.29	0.44	0.48
Cronbach's Alpha	0.33	0.57	0.58

The first row in Table 2 shows the reliability coefficients estimated by the split-half method created by Guttman (1945) and Flanagan (1942). This method involves the use of the

difference in test scores between the two halves of the tests and yields unique estimation of test reliability. The second row of Table 2 gives a summary of reliability coefficients computed through the Spearman-Brown prophecy formula. Since the correlation coefficient obtained from two subtest halves may be an underestimate of the reliability coefficient for the full-length subtest, the Spearman-Brown prophecy formula provides corrected reliability coefficients. The third row in Table 2 indicates the reliability coefficients obtained by using Cronbach's alpha of computing the ratio of the sum of the item covariances to the total observed score variance. These Cronbach's alpha coefficients may also be obtained by using the Kuder-Richardson formulas (KR-20 and KR-21) or Hoyt's method (1941). As can be seen from Table 2, the first section of the IERT, Vocabulary, has rather low reliability coefficients. The low reliability coefficients are a suggestion that the eight items in the Vocabulary do not "hold together" too well. In other words, the internal consistency of the Vocabulary Section is low. In the next section of this paper, *Estimated Validity*, we will examine the items in details and determine which items are not holding well with the common factor, vocabulary. From Table 2, the second section of the test, Comprehension, has a higher reliability than Vocabulary. Given the fact that there are only 7 items in this section, the relatively low reliability coefficients

may be regarded as acceptable, on the basis that if more items were put into the section the reliability coefficients would be higher. Also from Table 2, the reliability coefficients for the total test are not as high as they should be. This is mainly due to the fact that Vocabulary has low reliability coefficients which in turn dragged down the reliability of the whole IERT. Nevertheless, since alpha coefficients are the lower bound of the proportion of variance in the test scores explained by common factors underlying item performance, the actual internal consistency of these two subtests as well as the total test of the IERT may be higher than the estimated reliability coefficients in Table 2.

With the reliability coefficient alpha for the total Intermediate English Reading Test reported in the previous section, the standard error of measurement of the IERT was estimated to be 1.78. The standard error of measurement was derived from averaging the individual error standard deviations.

ESTIMATED VALIDITY

Three types of validation procedures were used for the Intermediate English Reading Test: content validation, criterion-related validation, and construct validation.

In conducting the content validation studies, our purpose was to assess whether the items in the IERT adequately represented the two performance domains of English at the intermediate level. In other words, we were interested not in whether the testees were able to answer the particular questions in the IERT, but rather in their abilities to answer similar questions in the tested domains. To do this, expert ESL teachers were invited to examine the IERT. All the items of the IERT were found fitting the test objectives, although some suggestions for test content and layout were made.

Criterion-related validation of the IERT was carried out in the form of concurrent validity study. Since the pilot testees were from two categories of intermediate classes, lower intermediate and upper intermediate, placed by the existing departmental Intermediate English Tests[5], a one-way analysis of variance (ANOVA) was run to see if the IERT results were consistent to those from the placement test. The rationale here was that, if the IERT was a valid intermediate English test, the IERT scores of the lower-intermediate English students would be significantly lower than the IERT scores of the upper-intermediate English students. The ANOVA results showed that the mean IERT score of the lower-intermediate English testees was 6.51 and that the mean IERT score of the upper-intermediate English testees was 8.86. The difference between

the two means is statistically significant, $F(1, 341)=10.767$, $p < 0.001$, confirming the validity of the IERT as a whole.

In an effort to determine how well the two parts of the IERT discriminate the testees, ANOVA was also run on each of the two parts of the IERT. The ANOVA results showed that the lower-intermediate group had a mean score of 3.17 and that the mean score of the upper-intermediate group was 4.79 in the first part of IERT, Vocabulary. The difference between the two means is passingly significant, $F(1, 341)=3.94$, $p<0.051$. This barely passing significance of the F ratio further indicates that Vocabulary may need some more work on it.

To investigate where the problem lies, the eight items in Vocabulary were correlated with each other and with the subtotal score of the Vocabulary. Table 3 shows the correlations among these eight items and their correlations with the subtotal score in this section of the test.

Table III.

Correlation matrix of the eight items

in the Vocabulary of IERT

	1	2	3	4	5	6	7	8	Total
Item1	-1.0000								0.2200
Item2	0.1766	1.0000							0.2069
Item3	0.0897	0.0682	1.0000						0.2082
Item4	0.1216	0.0802	0.0125	1.0000					0.0927
Item5	0.1208	0.0163	0.0583	0.0193	1.0000				0.2084
Item6	0.0411	0.1289	-.0436	0.0861	-.0185	1.0000			0.0975
Item7	0.0138	0.1167	0.0376	-.0462	0.0796	0.1091	1.0000		0.2071
Item8	0.1000	0.0283	0.1128	0.0157	0.0561	-.0005	0.0260	1.0000	0.2089

From the table, it appears that two items, Items 4 and 6, need some more work. These two items have low or negative correlations with the rest of the items in this section of the test. They also have the lowest correlations with the subtotal score of the section. This means that these two items do not cling well with the rest of the items in this section and therefore cannot be assumed to measure the same latent trait as the rest of the items do.

The second part of the IERT, Comprehension, seems to have a better validity than the first part. The lower-intermediate group had a mean score of 3.87, and the mean score for the upper-intermediate group was 4.60. ANOVA results suggested that the difference between the two means are statistically significant, $F(1, 341)=15.082$, $p<0.000$. The correlations among the seven items and their correlations with the subtotal score of the section are higher than those in the first section, Vocabulary. Table 4 is a summary of these correlations:

TABLE IV.

Correlation matrix of the seven items

in the Comprehension of the IERT

	9	10	11	12	13	14	15	Total
Item9	1.0000							0.3129
Item10	0.2827	1.0000						0.4147
Item11	0.1602	0.1603	1.0000					0.3455
Item12	0.2057	0.3304	0.1111	1.0000				0.3298
Item13	0.1472	0.1077	0.0834	0.1694	1.0000			0.2855
Item14	0.1554	0.0980	0.1095	0.1252	0.1621	1.0000		0.3574
Item15	0.0775	0.0731	0.1949	0.1381	0.2626	0.2003	1.0000	0.2934

The first subtest of the IERT has a correlation of 0.378 with the total IERT; the second subtest, 0.418. The correlation between these two subtests is 0.393.

For construct validity study, factor analysis was used. To reflect the complexity of the real world, latent factors were assumed to be correlated and thus quartimax (Jennrich and Sampson, 1966) was chosen as the rotation method. The results of the factor analysis, generally speaking, lent support to the test structure of the IERT. Five factors were extracted out of the 15 items. These 5 factors accounted for 71 percent of the total variance of the IERT. No unique factors were found. Except for one item, all the items formed clustered to reflect the construct structure of the IERT. The first two factors basically account for the first section, Vocabulary, of the IERT, and the last three factors are responsible for the second section, Comprehension. Table 5 shows the loadings of the 15 items on the 5 factors.

TABLE V.

Factor loadings of the 15 IERT items

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5
ITEM10	**.72591**	.04595	.24764	-.03455	.02495
ITEM12	**.70314**	.02568	-.15146	.13952	.00669
ITEM9	**.51643**	.34225	.08676	-.18142	-.18037
ITEM8	-.04231	**.67383**	.05945	-.14887	.14092
ITEM15	.02493	**.64081**	-.24401	.13388	.14661
ITEM14	.16465	**.57395**	-.08154	.03888	-.20581
ITEM11	.21810	**.35698**	.08171	.24937	.05655
ITEM4	.11019	-.13053	**.65654**	-.00835	.06343
ITEM13	.32721	.28729	**-.51573**	.03639	.14544
ITEM1	.35168	.23855	**.45386**	.08319	.13646
ITEM7	.03730	-.05067	-.15856	**.75989**	.13769
ITEM2	-.01675	.30109	.37716	**.49199**	-.03984
ITEM6	.06393	.07388	.15942	**.48540**	-.43713
ITEM3	-.07731	.17894	.22057	.00912	**.70648**
ITEM5	.37773	-.02433	-.19108	.17076	**.49542**

CONCLUSION

Generally speaking, the IERT may be regarded as a reliable reading test for people whose English is at the intermediate level. The relatively moderate reliability of the IERT is acceptable, because after all IERT is only one subtest of a battery of intermediate English tests. When used in conjunction with the other tests of the whole Intermediate English Desk Test (paragraph writing, grammar and oral interview), it is believed that the IERT will provide valid information on testees' English proficiency in reading. Nonetheless, further follow-up studies should be made to monitor the reliability and validity of the IERT, particularly, its first section, Vocabulary, which seems to have low reliability as suggested by the pilot testing results.

NOTES

1. We would like to express our thanks to Barbara Breen, Patricia Mori, Mary Sullivan, Jane McCourt, and Sheila Singh for their contributions to this report. They helped us with test administration and data collection.

2. ELA (English Language Assessment) is a battery of English proficiency tests developed by Vancouver Community College and used by a number of post-secondary educational institutions in British Columbia for admission and placement of students whose first language is other than English.

3. Calculated with the formula: 3(mean - median)/standard deviation.

4. Calculated with the formula: quantile deviation/(90th percentile-10th percentile).

5. The Intermediate English Test is a battery of English proficiency tests for intermediate English learners in the English Language Skills Department of KEC and is copyrighted by Vancouver Community College.

REFERENCES

1. Crocker, Linda & Algina, J. *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winston, 1986.

2. Ebel, R. L. *Measuring Educational Achievement*. Englewood Cliffs, N.J.: Prentice-Hall, 1965.

3. Flanagan, J.C. The critical incident technique. *Psychological Bulletin*, 51 (1954), 327-358.

4. Guttman, L. A basis for analyzing test-re-test reliability. *Psychometrika*, 10 (1945), 255-282.

5. Harman, H. H. *Modern Factor Analysis* (2nd Ed.). Chicago: University of Chicago Press, 1967.

6. Hoyt, C. J. Test reliability estimated by analysis of variance. *Psychometrika*, 6 (1941), 153-160.

7. Jennrich, R. I., and Sampson, P. F. Rotation for simple loadings. *Psychometrika*, 31 (1966), 313-323.

8. Kuder, G. F., and Richardson, M. W. The theory of the estimation of test reliability. *Psychometrika*, 2 (1937), 1511-160.

9. Rulon, P. J. A simplified procedure for determining the reliability of a test by split-haves. *Harvard Education Review*, 9 (1939), 99-103.

Intermediate English Reading Test

Please read the following passage and answer the questions on the next two pages.

Raccoons

Fences usually keep dogs out but not cats. However nowadays, in both small towns and cities, fences are not helpful in keeping another cat-like animal away either. This animal is an excellent climber because it has very muscular hind legs and a strong tail. Perhaps you have seen this animal at night and thought it was a big cat. But it was not! It was a raccoon, an animal which is becoming a pest all over North America.

Raccoons can live practically anywhere as long as they can find water, food and a dry shelter to sleep during the day. They can eat and will eat almost anything from the fish and plants in a person's pond to the contents in one's garbage can. Raccoons have very skillful front paws with five fingers that can grasp and hold things very easily.

These animals are not true hibernators like bears, which sleep all winter. Their body temperature drops only slightly and they wake up very easily. Although there may be snow on the ground, they go in search of food. They have a double fur coat which protects them from the cold, wind, rain and snow.

Raccoons are intelligent, curious and have an excellent memory. They look adorable but they are wild animals. They often attack and kill cats. At times these ring-tailed, bright-eyed creatures look straight at a person and seem to say, "I dare you to chase me away!" Raccoons have really become a nuisance in many North American cities.

Vocabulary - Choose the word from the list on the right that is closest in meaning to the word on the left.

1. muscular (line 5)
   a. strong
   b. skillful
   c. heavy
   d. large

2. hind (line 5)
   a. long
   b. short
   c. front
   d. back

3. pest (line 8)
   a. danger
   b. killer
   c. nuisance
   d. pet

4. grasp (line 14)
   a. put down
   b. feel
   c. take hold of
   d. break

5. hibernators (line 16)
   a. animals that hunt in winter
   b. animals that live in caves
   c. animals that sleep all winter
   d. animals that eat only meat

6. curious (line 22)
   a. patient
   b. dangerous
   c. inquisitive
   d. tame

7. adorable (line 23)
   a. wild
   b. lovable
   c. tame
   d. safe

8. nuisance (line 27)
   a. advantage
   b. attraction
   c. problem
   d. disaster

Reading Comprehension Questions

9. Raccoons look like

   a. cats.
   b. dogs.
   c. pests.
   d. bears.

10. A fence will not keep out a raccoon because

    a. the raccoon will break it down.
    b. the raccoon will make a hole in it.
    c. the raccoon will climb over it.
    d. the raccoon will jump over it.

11. Raccoons

    a. prefer to eat fish and plants.
    b. only eat garbage.
    c. usually eat cats.
    d. will eat almost anything.

12. Raccoons

    a. sometimes go out to look for food during the winter.
    b. sleep through the whole winter.
    c. never go out in the cold weather.
    d. have no protection against the cold weather.

13. Which of the following sentences is NOT true?

    a. Raccoons use their front paws like hands.
    b. Raccoons are afraid of people.
    c. Raccoons can be dangerous for house pets.
    d. Raccoons are usually seen at night.

14. In the last paragraph, the writer

    a. says that raccoons cannot be chased away.
    b. suggests that raccoons can be dangerous.
    c. thinks that raccoons would make good pets.
    d. suggests that raccoons need to live in cities.

15. The MAIN IDEA of this reading passage is

    a. fences don't keep out raccoons.
    b. raccoons sometimes eat garbage.
    c. raccoons are troublemakers in many places.
    d. wild animals can live in cities.